Bayesian stock assessment using catch-age data and the sampling – importance resampling algorithm

Murdoch K. McAllister and James N. Ianelli

Abstract: A Bayesian approach to fisheries stock assessment is desirable because it yields a probability density function (pdf) of population model parameters. This pdf can help to provide advice to fishery managers about the consequences of alternative harvest policies and convey uncertainty about quantities of interest such as population biomass. In stock assessment, catch-age data are commonly used to estimate population parameters. However, there are few catch-age analyses that use Bayesian methods. In this paper, we extend the sampling – importance resampling algorithm so that a pdf of population model parameters can be estimated using catch-age data and indices of relative abundance. We illustrate the procedure by estimating a 54-parameter pdf for yellowfin sole (*Limanda aspera*) in the eastern Bering Sea. The example demonstrates how catch-age data can markedly improve Bayesian estimation, and also illustrates the potential for significant biases in treating trawl survey abundance indices as absolute measures of stock size.

Résumé : Une approche bayésienne à l'évaluation des stocks de poissons exploités par la pêche est souhaitable parce qu'elle donne une densité de probabilité (dp) des paramètres du modèle de population. Cette dp peut aider les gestionnaires des pêches à cerner les conséquences d'autres politiques de récolte et véhiculer l'incertitude au sujet des quantités d'intérêt, comme la biomasse de la population. Toutefois, il existe peu d'analyses des prises par âge qui font appel aux méthodes de Bayes. Dans le présent article, nous étendons l'algorithme échantillonnage – rééchantillonnage d'importance de sorte que l'on puisse estimer une dp des paramètres du modèle de population en utilisant les données sur les prises par âge et les indices d'abondance relative. Nous illustrons la méthode en estimant une dp à 54 paramètres pour la limande à nageoires jaunes (*Limanda aspera*) dans l'est de la mer de Bering. L'exemple démontre comment les données sur les prises par âge peuvent servir à améliorer grandement l'estimation bayésienne et, également, illustre la possibilité d'introduire des biais significatifs lorsqu'on traite les indices d'abondance obtenus à l'aide de relevés par chalutage comme étant des mesures absolues de la taille des stocks.
[Traduit par la Rédaction]

Introduction

Many of the models used to provide advice to fishery managers, for example, about total allowable catches (TACs), require fixed values for input parameters (e.g., natural mortality rate (M), recruitment, and fishery selectivity). These assumptions often mask the level of uncertainty inherent in stock assessments. Bayesian statistical approaches have recently been applied to incorporate such uncertainties in advice to fishery managers (Punt 1993; McAllister et al. 1994; Stocker et al. 1994; Walters and Ludwig 1994; Walters and Punt 1994; Ianelli and Heifetz 1995; Raftery et al. 1995; Kinas 1996). The resulting Bayesian probability distributions can readily convey to fishery managers the uncertainty in derived model quantities such as the biomass of a fish population. Bayesian

M.K. McAllister.[1] Fisheries Research Institute, Box 357890, University of Washington, Seattle, WA 98195-7980, U.S.A.
J.N. Ianelli. National Marine Fisheries Service, Alaska Fisheries Science Center, 7600 Sand Point Way NE, Seattle, WA 98115, U.S.A.

[1] Present address: Renewable Resources Assessment Group, Imperial College, 8 Princes Gardens, London, SW7 1NA, U.K.

methods can also help managers to assess the likelihood of a variety of outcomes, e.g., rates of population decline, that could result from a proposed management action such as setting a harvest level or TAC.

A drawback of adapting Bayesian methods to large (many parameter) nonlinear models such as those used for age-structured fisheries stock assessment has been with calculating the multidimensional integrals that are fundamental to the joint probability distribution of model parameters. Monte Carlo methods are the most common methods for approximating such high-dimensional integrals. These include Markov Chain Monte Carlo methods (e.g., Metropolis–Hastings algorithm) and importance sampling (e.g., sampling – importance resampling (SIR) algorithm) (Metropolis et al. 1953; Hastings 1970; Berger 1985; Rubin 1987, 1988; Gelfand and Smith 1990; Smith 1991; West 1993). We first provide a brief overview of the application of these methods in stock assessment. We then suggest an extension of one of these methods so that catch-age data can be used in addition to relative abundance data.

Of the Monte Carlo methods, the SIR algorithm (Rubin 1987, 1988) has been applied most often in Bayesian fisheries assessments (e.g., Punt 1993; McAllister et al. 1994; Stocker et al. 1994; Raftery et al. 1995). In a study that was not explicitly Bayesian, Francis et al. (1992) used a method analogous to the SIR algorithm to calculate the biological

risks of alternative harvesting policies for orange roughy (*Hoplostethus atlanticus*). Only two of several types of the population model input parameters used (i.e., average unfished biomass, B_0, and recruitment anomalies, ϵ_y) were treated as uncertain. Raftery et al. (1995) used the SIR algorithm to estimate probability distributions for bowhead whale (*Balaena mysticetus*) population parameters. The procedure used considerable prior information on input parameters for a deterministic population dynamics model and estimated quantities relating to the current status of the population such as the replacement yield in the current year. McAllister et al. (1994) generalized these approaches to provide the flexibility to incorporate uncertainty in recruitment deviates about a stock–recruit function, informative priors for several model parameters, and a Bayesian decision analysis of alternative harvesting policies. They estimated the posterior by fitting an age-structured population dynamics model to relative indices of abundance and using age-structured data to construct priors on recruitment anamolies, ϵ_y. However, there is often relatively little information about input parameters to population dynamics models in relative abundance indices. In contrast, the incorporation of age frequency or catch-age data from the commercial catch and (or) from research surveys to the likelihood function can help to improve statistical inference about input parameters and derived quantities such as stock biomass.

In this paper, we extend the use of the SIR algorithm so that a population dynamics model can be fitted to fishery catch-age data and survey data. We illustrate the algorithm using eastern Bering Sea yellowfin sole (*Limanda aspera*) data. In the example, we present a new procedure to construct a prior probability distribution for the constant of proportionality for trawl survey abundance indices that uses survey data and expert judgment. The example illustrates the potential improvements in statistical inference that can be obtained by including catch-age data in a Bayesian stock assessment. The example also illustrates the large biases that can result from using area-swept survey abundance indices as absolute rather than relative measures of stock size.

A conceptual framework for a Bayesian stock assessment

McAllister et al. (1994) defined a conceptual framework for a Bayesian stock assessment: a population dynamics model is fitted to data to construct a posterior probability density function (pdf) of the vector of model parameters, θ, (e.g., B_0, M, ϵ_y). A "state of nature" is one potential realization of the set of all possible values for θ, say $θ_i$. The probability that a given state of nature $θ_i$ is true conditioned on the data, $P(θ_i | \text{data})$, is given by

$$(1) \qquad P(θ_i | \text{data}) = \frac{L(\text{data} | θ_i)p(θ_i)dθ}{\int L(\text{data} | θ)p(θ)dθ},$$

where $p(θ_i)$ is the prior probability density for state of nature i and $L(\text{data} | θ_i)$ is the likelihood function for the data evaluated at $θ_i$. The value $p(θ)$ represents the probability that a given state of nature is true prior to obtaining a set of data that can further our ability to discriminate among alternative $θ_i$.

Using sampling – importance resampling and catch-age data to construct probability distributions for quantities of interest

A sample from the posterior distribution, $P(θ | \text{data})$, can be used to estimate marginal probability distributions for quantities of interest, such as current population biomass, that are known functions of the parameters, θ. This sample can be drawn using the SIR algorithm where it is possible to generate a sequence of i.i.d. (independent and identically distributed) random variables $(θ_1, θ_2, \ldots, θ_m)$ that have a common probability density function, $h(θ)$, called an importance function (Berger 1985). McAllister et al. (1994) summarized the statistical basis for the SIR algorithm. The procedure requires that for each vector $θ_k$ that is drawn from $h(θ)$, an importance ratio is calculated:

$$(2) \qquad w(θ_k) = \frac{L(\text{data} | θ_k)p(θ_k)}{h(θ_k)}.$$

The assignment of importance ratios forms a discrete distribution $F(θ | \text{data})$ over $(θ_1, θ_2, \ldots, θ_m)$ placing mass

$$(3) \qquad F(θ_k | \text{data}) = \frac{w(θ_k)}{\sum_{k=1}^{m} w(θ_k)}$$

on each $θ_k$. The distribution, $F(θ | \text{data})$ over $(θ_1, θ_2, \ldots, θ_m)$, approximates the actual posterior distribution, $P(θ | \text{data})$. This approximation improves as m increases (Berger 1985, p. 263; McAllister et al. 1994).

To estimate the marginal posterior distribution for a quantity of interest that is a known function of θ (e.g., $g(θ)$), the discrete distribution $F(θ | \text{data})$ over $(θ_1, θ_2, \ldots, θ_m)$ could be used (i.e., each $g(θ_k)$ would be assigned the mass $F(θ_k | \text{data})$ over $(θ_1, θ_2, \ldots, θ_m))$ (Berger 1985, p. 263). However, this may be numerically inefficient because m is often very large. Therefore, a random sample of size n, with $n < m$, is taken from $F(θ | \text{data})$ over $(θ_1, θ_2, \ldots, θ_m)$ (i.e., the probability of drawing $θ_1, θ_2, \ldots, θ_m$ is proportional to $F(θ_1 | \text{data})$, $F(θ_2 | \text{data}), \ldots, F(θ_m | \text{data}))$. This step is the resampling part of the SIR algorithm. The empirical distribution for $g(θ)$ is then given by $g(θ_1), g(θ_2), \ldots, g(θ_n)$ with mass $1/n$ placed on each of the $g(θ)$s.

A key task of importance sampling is that of finding a suitable $h(θ)$ (van Dijk and Kloek 1983; Berger 1985). This task becomes increasingly challenging as the number of parameters increases, and the posterior becomes more complex (e.g., multimodal). Oh and Berger (1992) suggested the following three guidelines for selecting an importance function. (*i*) It should be easy to generate draws from. (*ii*) The tails of $h(θ)$ (the value $h(θ)$ at extreme values of θ) should not be less dense than the tails of the posterior; otherwise, estimated posterior expectations of quantities of interest may have large variance or even fail to converge with increasing m. (*iii*) The density function $h(θ)$ should mimic the true posterior density function reasonably well (e.g., in skewness and covariance). If the data are not very informative, the model is fairly simple (e.g., deterministic), or adaptive importance sampling is used (see Discussion), the prior pdf can sometimes serve as an adequate importance function (Francis et

$θ_i$.

al. 1992; Kinas 1993; Punt 1993; McAllister et al. 1994; Raftery et al. 1995). However, in many instances, such as when catch-age data are available, the prior will be an inefficient importance function.

The multivariate Student t distribution, $t(\theta\,|\,d,\,\mu,\,V)$, has often been used as an importance function (van Dijk and Kloek 1983; Geweke 1989; West 1992; Kinas 1993):

$$(4) \qquad t(\theta\,|\,d,\mu,V) = \frac{\Gamma((d+p)/2)}{(\det V)^{1/2}(\pi d)^{p/2}\Gamma(d/2)}$$

$$\times \left(1 + \frac{1}{d}(\theta - \mu)'V^{-1}(\theta - \mu)\right)^{-(d+p)/2},$$

where d is the degrees of freedom ($d > 0$), μ is the mean if $d > 1$, and V is a positive definite $p \times p$ matrix $(\text{cov}(\theta) = dV/(d - 2)$ if $d > 2)$ (from Berger 1985).

This distribution can be useful provided that each parameter in θ is defined on the interval $-\infty$ to ∞. An advantage of this distribution over the multivariate normal is that the tails of the distribution can be adjusted (by changing d) so that the density is slightly greater than the density in the tails of the posterior distribution. Furthermore, the mean, μ, can be estimated by the mode of the posterior by nonlinear minimization. V can be estimated by the negative inverse of the estimated Hessian matrix, i.e., the $p \times p$ matrix of second derivatives $(\partial^2 \log(P(\theta\,|\,\text{data})))/(\partial\theta^i\partial\theta^j)$ where θ^i and θ^j are parameters i and j in θ, evaluated at the mode of $P(\theta\,|\,\text{data})$ (van Dijk and Kloek 1983).

Steps for developing an importance function for use with age data

The following steps can be used to develop an importance function using the multivariate t density.

(1) Construct an informative prior for key model parameters such as natural mortality rate, M, and the coefficient of variation (CV) for recruitment variability, σ_r (see example below and Appendix 3).

(2) Construct likelihood functions for each of the types of data (see below).

(3) Use nonlinear minimization to find the posterior modal values for θ, $\theta_{P(\theta\,|\,\text{data})}$. This is achieved by finding the values in the vector θ that minimize the quantity $-(\log(L(\text{data}\,|\,\theta)) + \log(p(\theta)))$.

(4) Estimate the covariance, V, about the mode of the joint posterior by estimating the negative of the inverse of the Hessian matrix, $\hat{V}_{P(\theta\,|\,\text{data})}$ (Arnold 1990).

(5) Set the importance function to $t(\theta\,|\,d,\,\hat{\theta}_{P(\theta\,|\,\text{data})},\,\hat{V}_{P(\theta\,|\,\text{data})})$ where d is set to some value between 15 and 25 (values in this range have provided satisfactory results).

(6) Where appropriate, transform the parameters so that the importance function approximates the skewness in the posterior (van Dijk and Kloek 1983). If a transformation is used, the multivariate t density must be multiplied by the determinant of the Jacobian (van Dijk and Kloek 1983; Arnold 1990) to obtain the density function for the untransformed parameters, $h(\theta)$ (see example below and Appendix 1). This is necessary for calculating $w(\theta_k)$, since the terms of $w(\theta_k)$ are in θ, not in the transformed θ (eq. 2).

(7) Follow the standard steps for importance sampling (see

below). This requires that parameter values be sampled or drawn from the importance function. We used the Choleski decomposition of the variance–covariance matrix to create parameter draws (Ripley 1987; Nash 1990; Kinas 1993).

(8) Modify the importance function if sampling from it appears to be too inefficient for estimating the posterior. A crude index of sampling efficiency is the maximum importance ratio, $w(\theta_k)$, taken as a proportion of the sum of all $w(\theta_k)$ over all of the draws. If the sample was taken from the posterior, this proportion would be equal to $1/m$, since each $w(\theta_k)$ would be equal to some constant. A target, for example, might be a maximum $w(\theta_k)$ of less than 5% of the total. However, this target may vary depending on the situation.

It is advisable to run replicate importance sampling trials with different random number sequences to evaluate whether the target can be consistently obtained and whether the replicate posteriors are sufficiently similar. In the example below, a maximum $w(\theta_k)$ of 4% of the total appeared to provide estimates of posterior pdfs sufficiently precise for stock assessment and decision analysis (see Results).

Another indicator of sampling efficiency is the CV in $w(\theta_k)$ where the CV is given by

$$(5) \qquad \text{CV}(w(\theta_k)) = \frac{\text{SD}(w(\theta_k))}{\frac{1}{m}\sum_{k=1}^{m} w(\theta_k)}$$

In the example below, increases in CVs to above 60 after about 50 000 draws of θ from $h(\theta)$ indicated that the importance function was inefficient and required adjustment. It is also useful to keep track of the parameter values in draws with the maximum $w(\theta_k)$ to compare with the modal estimates (Oh and Berger 1992). If the drawn values with maximum $w(\theta_k)$ are in the tails of the posterior distribution, the tails of the importance function may be too thin and need to be modified. This can be remedied by reducing the degrees of freedom, deemphasizing priors in estimating the Hessian matrix (see example below), or by reparameterizing to adjust for skewness. The split t distribution has been suggested as a more flexible distribution than the multivariate t for making adjustments for skewness (Geweke 1989). Importance sampling efficiency can also be improved by using estimates of the covariance matrix of the posterior distribution obtained from the first round of importance sampling in place of the Hessian estimate (van Dijk and Kloek 1983). This is because the estimate of posterior covariance provided by the Hessian matrix is a "local" estimate of covariance at the mode of the posterior; the alternative is a "global" estimate of the posterior covariance (van Dijk and Kloek 1983).

Steps of the SIR algorithm when age data are used

Below, we summarize the steps required to extend the SIR algorithm in McAllister et al. (1994) to the use of catch-age data from commercial catches and research surveys.

(1) Once an importance function, $h(\theta)$, has been chosen, draw one set of values θ_k for the model inputs from $h(\theta)$ (k refers to draw number k, since this step is repeated many times). For example, draw one set of values for B_0, α, M, q, s_{50}, a_{50}, and the recruitment residuals, ϵ_y, from $h(\theta)$ where α is a stock–recruit parameter, q is a catchability coefficient for

trawl survey estimates of biomass, s_{50} and a_{50} are selectivity parameters, and the other parameters have been defined above (see below).

(2) Project the population dynamics model from the initial year of the fishery forward to the current year (e.g., see below), using the values of model inputs, θ_k, drawn in step 2 and the known catches to predict stock biomass and catch-age data. Note that it is unnecessary to assume that unfished biomass is at equilibrium. Recruitment is assumed to be lognormally distributed about the assumed (e.g., Ricker) stock–recruit function (eq. 9). Values for all other parameters are assumed to be stationary over time. (The projection of the model from unfished conditions up to the current year will be referred to as historic trajectory k.)

(3) Evaluate the likelihood function, $L(\theta_k | \text{data})$, mass of the prior, $p(\theta_k)$, and mass of the importance function, $h(\theta_k)$, for the kth draw from $h(\theta)$, i.e., historic trajectory k (e.g., see below).

(4) If any of the observed catches exceeds the predicted exploitable biomass, set the prior to 0. This can be achieved by adding a penalty function to the log likelihood (see eqs. 20 and 21).

(5) Calculate the importance ratio, $w(\theta_k)$, for historic trajectory k (eq. 2).

(6) Repeat steps 2–6 until the desired level of sampling precision is obtained. In the example below, $m = 3\ 000\ 000$ reduced the maximum $w(\theta_k)$ to 4% of the total. With a fast desktop computer (e.g., a Pentium 100) and programming language (e.g., C), this could potentially be achieved in an overnight run.

(7) Draw a large number (e.g., $n = 10\ 000$) of θ_k randomly with replacement from $(\theta_1, \theta_2, \ldots, \theta_m)$ with probabilities proportional to $w(\theta_1), w(\theta_2), \ldots, w(\theta_m)$. The resulting sample $(\theta_1, \theta_2, \ldots, \theta_n)$ approximates an i.i.d. sample from the joint posterior pdf of θ. This sample could be used when performing the decision analysis (McAllister and Pikitch 1997). It is also useful for estimating the marginal posterior probabilities and covariances for model input parameters (McAllister et al. 1994).

Estimating a joint posterior probability distribution of population dynamics model input parameters for yellowfin sole in the eastern Bering Sea

The exploitation history of yellowfin sole in the eastern Bering Sea extends back to the mid 1950s. In the early 1960s, the population was heavily exploited by foreign trawl vessels with annual catches up to about 550 000 t (Wilderbuer et al. 1992) (Fig. 1). Since the early 1970s, fishery removals have dropped off to about 100 000 t annually and recruited biomass appears to have recovered to levels over 2 million t (Wilderbuer et al. 1992). The current U.S. domestic fishery is managed by TAC quotas that are set annually. Assessment scientists provide the estimates of stock biomass and allowable biological catch (ABC). The ABC is obtained by applying to the catch equation the estimate of fishing mortality rate that will reduce spawner biomass per recruit to 35% of that under unexploited conditions, $F_{35\%}$ (Clark 1991). The assessment typically uses age-structured data from the commercial fishery and a trawl survey and, also, absolute

Fig. 1. Catch biomass (t), trawl survey (t/10), and TAC (t) series for yellowfin sole in the eastern Bering Sea.
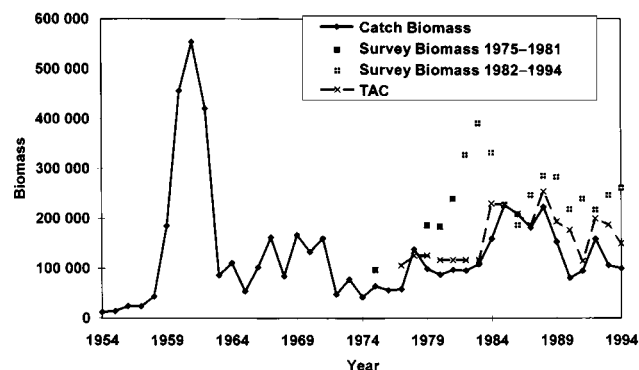
estimates of biomass from the trawl survey. However, controversy exists over abundance estimates because recent gear efficiency studies (D. Somerton, Alaska Fisheries Science Center (AFSC), National Marine Fisheries Service (NMFS), Seattle, Wash.) indicate that, for some species, trawl survey gear is considerably more efficient than is commonly assumed. Below, we describe the model that we used to assess the status of yellowfin sole and the specifications used to estimate a joint posterior probability distribution for the model parameters and other quantities of interest for fisheries management.

Population dynamics model

The model is age-structured and assumes that annual recruitment is lognormally distributed about a Ricker stock–recruit function. The model assumes that recruitment to the fishery follows a logistic function of age such that a fixed fraction of individuals within an age class will recruit annually to the fishery from individuals not previously vulnerable to fishing. In each year, fishing mortality is assumed to occur after natural morality has occurred. In 1954, the initial year of the fishery, population biomass is assumed to be distributed about B_0 and age structure is partially stationary ($\epsilon_y = 0$ for ages >8 years).

Resource dynamics

The dynamics of animals of age 2 and above are governed by the equation

$$N_{y+1,a} = e^{-M}(N_{y,a-1}(1 - H_y) + U_{y,a-1}\tau_a)$$

$$2 \leq a < n$$

(6)
$$U_{y+1,a} = e^{-M}U_{y,a-1}(1 - \tau_a) \qquad 2 \leq a < m$$

$$N_{y+1,m} = e^{-M}(N_{y,m} + N_{y,m-1})(1 - H_y) \qquad a = m$$

where $N_{y,a}$ is the number of recuited animals of age a at the start of the year y, $U_{y,a}$ is the number of unrecruited animals of age a at the start of year y, τ_a is the proportion of unrecruited animals of age $a - 1$ that recruit at age a:

(7) $\qquad \tau_a = (\phi_a^f - \phi_{a-1}^f)/(1 - \phi_{a-1}^f),$

ϕ_a^f is the fraction of animals of age a that would be recruited

if the population were at its deterministic unexploited equilibrium level:

$$(8) \qquad \phi_a^f = \frac{1}{1 + \exp(-s_{50}^f(a - a_{50}^f))},$$

s_{50}^f is the slope of the fishery recruitment ogive at 50% recruited, a_{50}^f is the age at 50% recruited, M is the instantaneous rate of natural mortality on animals, H_y is the harvest rate during year y and occurs after natural mortality (all recruited animals are assumed to be equally vulnerable to fishing), and m is the maximum (lumped) age-class (all animals in this and the previous age-class are recruited and mature).

Births

$$(9) \qquad \begin{array}{l} N_{y,1} = 0 \\[4pt] U_{y,1} = (E_{y-1} \exp(\alpha_0(1 - E_{y-1}/\beta)))e^{\epsilon_y - I_y \sigma_r^2/2}, \end{array}$$

where E_y is the total reproductive potential of the population during year y at the time of spawning:

$$(10) \qquad E_y = e^{-M} \sum_{a=1}^{m} \mu_a f_a (N_{y,a}(1 - 0.5H_y) + U_{y,a}),$$

μ_a is the fraction of mature fish at age a (see Table 3), f_a is the reproductive potential per individual fish of age a, and is assumed equal to w_a, w_a is the mass of a fish of age a (assumed to be constant throughout the year) (see Table 3):

$$(11) \qquad w_a = \delta_1(L_a)^{\delta_2},$$

$L_a = L_\infty(1 - e^{-k(a - t_0)})$, ϵ_y is the recruitment residual for year y: $e^{\epsilon_y - I_y \sigma_r^2/2}$ is referred to as the recruitment multiplier ($\epsilon_y \sim \text{Normal}(0, \sigma_r^2)$ (see Table 2), I_y is set to 1 if the model is used for decision analysis (McAllister and Pikitch 1997) and the year is greater than the last year in which ϵ_y is estimated (i.e., 1990) and 0 otherwise (setting I_y to 1 activates the lognormal correction error factor; thereby the expected value for recruits, $E[U_{y,1}]$, equates with the value given by the stock–recruit function), σ_r is the standard deviation of the log of the multiplicative fluctuations in births, and α_0, β are stock–recruit function parameters.

Initial conditions

The initial numbers-at-age (in 1954 for yellowfin sole) are given by the equation

$$(12) \qquad \begin{array}{ll} N_{y_1,a} = R_1 \phi_a \exp(-(a-1)M)e^{\epsilon_a} & 1 \leq a \leq m-1 \\[4pt] N_{y_1,m} = R_1 \exp(-(m-1)M)/ & \\ \qquad\quad (1 - \exp(-M)) & a = m \\[4pt] U_{y_1,a} = R_1(1 - \phi_a) & \\ \qquad\quad \times \exp(-(a-1)M)e^{\epsilon_a} & 1 \leq a \leq m-1 \end{array}$$

where R_1 is the number of 1 year olds at the deterministic equilibrium that corresponds to an absence of harvesting and ϵ_a is estimated only for cohorts that are included in the first year of age data. For yellowfin sole, catch age data start in 1964 and the plus group for the age data is 18 yr. Therefore,

the maximum age in which ϵ_a is estimated was age 8. For $a > 8$, ϵ_a was set to 0. A value for R_1 is calculated from the value for the virgin biomass at midyear (after natural mortality), B_0, using the equation

$$(13) \qquad R_1 = B_0 \bigg/ \bigg\{ \sum_{a=1}^{m-1} w_a \phi_a^f \exp(-aM) \\ + w_m \exp(-mM)/[1 - \exp(-M)] \bigg\}.$$

Values for the stock–recruit function parameters α and β are calculated from the values of R_1:

$$\alpha_0 = \alpha + 7M$$

$$\beta = \frac{\tilde{B}_0^S R_1}{(1 - \log(1/\tilde{B}_0^S)/\alpha_0)}$$

$$(14) \qquad \tilde{B}_0^S = \bigg(\sum_{a=1}^{m-1} \mu_a f_a e^{-aM} + \mu_a f_a$$

$$\times \exp(-mM)/(1 - \exp(-M)) \bigg),$$

where $\tilde{B}_0^S$ is the reproductive potential per recruit and α is the value for α that is selected from the prior distribution for α. α is given in units of individual age 7 recruits after natural mortality. The adjustment of α to α_0 is made to rescale the prior value to the units for recruits in the assessment model (in age 1 recruits before natural mortality).

Catches

The exploitation rate during year y, H_y, is obtained by solving the equation

$$(15) \qquad H_y = \frac{C_y}{B_y},$$

where C_y is the observed annual catch biomass during year y (assumed to be observed without error) and B_y is the exploitable biomass in year y:

$$(16) \qquad B_y = \sum_{a=1}^{m} w_a N_{y,a} e^{-M}.$$

Types of data used

The population dynamics model was fitted to two series of biomass estimates from the trawl surveys in the eastern Bering Sea: the first series is from 1975 to 1981 and the second from 1982 to 1994 (Table 1). The data were provided by G. Walters (AFSC, NMFS, Seattle, Wash.). The two series were treated separately because the trawl survey gear was changed in 1982 (see Gunderson 1993 for details). Both series were treated as indices of relative abundance, each with its own constant of proportionality. Both were assumed to have the same asymptotic selectivity function because the fit of the model to the data was not improved with separate selectivity estimates. The removals by survey gear were assumed to have an insignificant effect on stock size and to have occurred after recruitment and natural mortality and after half of the commercial catch had been removed.

Table 1. Total commercial catch biomass and trawl survey estimates of biomass of yellowfin sole in the eastern Bering Sea (tonnes).

Year	Catch	Survey	Year	Catch	Survey
1954	12 562	—	1975	64 690	972 500
1955	14 690	—	1976	56 221	—
1956	24 697	—	1977	58 373	—
1957	24 145	—	1978	138 433	—
1958	44 153	—	1979	99 017	1 866 500
1959	185 321	—	1980	87 391	1 842 400
1960	456 103	—	1981	97 301	2 394 700
1961	553 742	—	1982	95 712	3 275 300
1962	420 703	—	1983	108 385	3 910 600
1963	85 810	—	1984	159 526	3 320 300
1964	111 177	—	1985	227 107	2 277 400
1965	53 810	—	1986	208 597	1 866 400
1966	102 353	—	1987	181 428	2 465 800
1967	162 228	—	1988	223 156	2 854 600
1968	84 189	—	1989	153 170	2 832 180
1969	167 134	—	1990	80 584	2 183 800
1970	133 079	—	1991	95 000	2 393 300
1971	160 399	—	1992	159 038	2 172 900
1972	47 856	—	1993	106 101	2 465 400
1973	78 240	—	1994	100 000	2 610 481
1974	42 235	—	1995	100 000	—

Note: The CV in the survey biomass data was 0.19 (obtained using equations and data in McAllister 1995, Chap. 1).

The trawl survey data are assumed to be proportional to the biomass that is susceptible to the survey gear (i.e., B'_y):

$$(17) \quad B'_y = e^{-M} \sum_{a=1}^{m} \phi^t_a (N_{y,a}(1 - 0.5H_y) + U_{y,a}) w_a,$$

where ϕ^t_a is the proportion of individuals at age a that are vulnerable to the survey gear,

$$(18) \quad \phi^t_a = \frac{1}{1 + \exp(-s^t_{50}(a - a^t_{50}))},$$

s^t_{50} is the slope of the selectivity ogive at 50% selected, and a^t_{50} is the age at 50% selected.

Both series of abundance indices are treated as being log-normally as

$$(19) \quad \log(O^i_j) = \text{Normal}(\log(q_i E^i_j), \sigma^2_i),$$

where O^i_j is the jth value of the ith series, q_i is the constant of proportionality of the ith series, E^i_j is the model estimate of the quantity corresponding to O^i_j (e.g., B'_y in eq. 17), and α_i is the coefficient of variation for series i. The contribution of index i to the log likelihood is

$$(20) \quad \lambda^i = -0.5 n_i \log(2\pi) - \sum_j \log(\sigma_i O^i_j)$$
$$- \frac{0.5}{\sigma^2_i} \sum_j \log(O^i_j / (q_i E^i_j))^2,$$

where n_i is the number of data points for abundance index series i. The first two terms were ommitted from the calculations because they contain constants only. The values for the parameters in $q_i E^j_i$ are treated as unknown and estimated.

The proportions at age in the commercial catch are assumed to be given by

$$(21) \quad c^f_{a,y} = \frac{N_{a,y}}{\sum_{a'=1}^{m} N_{a',y}}$$

The proportions at age in the trawl survey are assumed to be given by

$$(22) \quad c^t_{a,y} = \frac{\phi^t_a((1 - 0.5H_y)N_{a,y} + U_{a,y})}{\sum_{a'=1}^{m} \phi^t_{a'}(1 - 0.5H_y)N_{a',y} + U_{a',y})}$$

In both the commercial catch-age data series (extending from 1964 to 1993) and the survey data series (1975 to 1993) (Table 1), the observed numbers at age in each year were assumed to be multinomially distributed. The contribution of series i to the log likelihood is

$$(23) \quad \lambda^i = \hat{T}^i \sum_{j=1}^{n_i} \sum_{a=1}^{m'} o^i_{a,j} \log(c^i_{a,j})$$

where $\hat{T}^i$ is the estimated effective sample size for the total number of fish sampled in each year, $o^i_{a,j}$ is the proportion at age from the original sample, $c^i_{a,j}$ is the predicted proportion of fish at age a in the jth year of series i, n_i is the number of years in series i, and m' is the plus group for the age data (18 yr for both series) (Appendix 2).

The number of fish that were effectively sampled at random in each age sample (i.e., effective sample size, Pennington and Vølstad 1994) was estimated to be 102 for the fishery series and 104 for the survey series (see Appendix 2). In the estimation, the values used in the multinomial likelihood function for each year were rounded to 100.

To fit an age-structured model to catch biomass data series, we assumed that the fraction of fish harvested in each recruited age class is equal to the ratio of the observed catch biomass to the predicted exploitable biomass and that the catch biomass series is known without error (Francis 1992; Francis et al. 1992; Punt et al. 1993; McAllister et al. 1994).

The log likelihood also incorporated a term for H_y that sets the prior equal to 0 when the harvest rate exceeds the maximum possible (similar to the likelihood function of the observed catch biomass values in Bence et al. 1993):

$$(24) \quad \lambda^5 = -0.5 n_5 \log(2\pi) - \sum_j \log(\sigma_5 C_j)$$
$$- \frac{0.5}{\sigma^2_5} \sum_j \log(C_j / \hat{C}_j)^2,$$

where σ_5 is the standard deviation in C_j (set very small, e.g., to 0.0002, and C_j are assumed to be measured without error), and $\hat{C}_j$ is given by:

$$(25) \quad \begin{array}{l} \hat{C}_j = C_j \text{ if } H_y < 0.99 \\ \hat{C}_j = 0.99 B'_y \text{ if } H_y \geq 0.99. \end{array}$$

Prior probabilities used

The procedure used a joint prior pdf of multiple model input parameters. For simplicity, the prior assumes that the parameters are all independent of each other. For example, prior to evaluating the data, no combinations of values for B_0 and M are assumed to be more or less credible than implied by the product of their respective prior pdfs. Thus, under the assumption of independence, prior pdfs are constructed individually for each parameter and the joint prior is the product of the priors for each parameter:

$$(26) \qquad p(\theta_k) = \prod_{i=1}^{54} p(\theta_k^i),$$

where $p(\theta_k^i)$ is the prior for parameter i in the vector θ_k.

The 54 estimated model input parameters included the following: (i) virgin biomass, B_0, (ii) stock productivity, α (reflecting the slope parameter of the Ricker stock–recruit function at 0 spawners, eq. 9), (iii) natural mortality rate, M, (iv) the constant of proportionality, which relates the trawl estimates of stock biomass to actual biomass, q_i, (v) the age at 50% recruited, a_{50}^f, (vi) the slope of the fishery recruitment ogive at 50% recruited, s_{50}^f, (vii) the age at 50% vulnerable to the trawl survey, a_{50}^t, (viii) the slope of the trawl selectivity ogive at 50% vulnerability to trawl survey gear, s_{50}^t, and (ix) the annual lognormal recuitment residuals ϵ_y from the years 1946 to 1990 (eq. 9). The values for ϵ_y for the years 1934–1945 were fixed at 0 to provide reasonably efficient estimates of the posterior (the first year of reported catches is 1954). This caused an underestimate of the posterior variance for quantities such as B_0. However, the Hessian estimates of the CV in the marginal posterior distributions were only slightly less than when ϵ_y for the years 1934–1990 were estimated (e.g., by 1.6% for B_0).

The procedure uses informative prior pdfs of parameters B_0, α, σ_r, M, and q_i (Table 2). These priors were constructed from stock–recruit data for other flatfishes and by consulting the literature and soliciting expert judgment. The prior for B_0 is uniformly distributed over the interval 500–5000 kt. The priors for α and σ_r were constructed using an approach in McAllister et al. (1994). The prior for α is Normal(1.52, 0.78²) and was obtained from stock and recruit data for 14 other populations of small-sized flatfishes (Fig. 2; and see Appendix J in McAllister 1995). The prior for σ_r is lognormal with a mean of 0.46 and an SD of 0.40 (Fig. 2; and see Appendix J in McAllister 1995). σ_r was fixed at the posterior modal estimate (0.40, obtained using nonlinear minimization) for estimating the posterior distribution with SIR because SIR was too inefficient when the prior for σ_r was used. McAllister et al. (1994) found that the estimate of the posterior was relatively insensitive to using an informative prior for σ_r versus the best estimate for the population. The prior pdf of M (Table 2) is lognormal and was parameterized by setting the 2.5 and 97.5 percentiles of the distribution to 0.08 and 0.18, a range consistent with knowledge of the life history of yellowfin sole, aging studies, and values used in stock assessments of 15 other populations of small-sized flatfishes (i.e., 0.10 to 0.12) (Wilderbuer et al. 1992; Myers et al. 1993).

The prior pdfs of q_i were lognormal with a mean of 0.44 and a CV of 0.59 for the first series and a mean of 0.97

Table 2. Priors of some population dynamics model parameters.

Parameter	Prior
B_0 (×1000 t)	Uniform(500, 5000)
α	Normal(1.52, 0.78²)
σ_r	lognormal(0.46, 0.40²)
ϵ_y ($y \in$ 1946–1990)	Normal(0, σ_r^2)
M(year⁻¹)	$M \sim$ lognormal(0.12, 0.2²)
q_{75-81}	lognormal(0.44, 0.59²)
q_{82-94}	lognormal(0.97, 0.51²)
a_{50}^f, a_{50}^t(years)	Normal(7.5, 3²), $a_{50}^f > 0$, $a_{50}^t > 0$
s_{50}^f, s_{50}^t	Normal(1.5, 2²), $s_{50}^f > 0$, $s_{50}^t > 0$

Fig. 2. (a) Empirical and fitted prior pdf of α in the Ricker stock–recruit function based on stock–recruit data for 14 flatfish stocks in Myers et al. (1993) (see Appendix J in McAllister 1995). (b) Empirical and fitted prior pdf of σ_r. (c) Plot of the modal estimate of σ_r versus α for the 14 flatfish stocks.
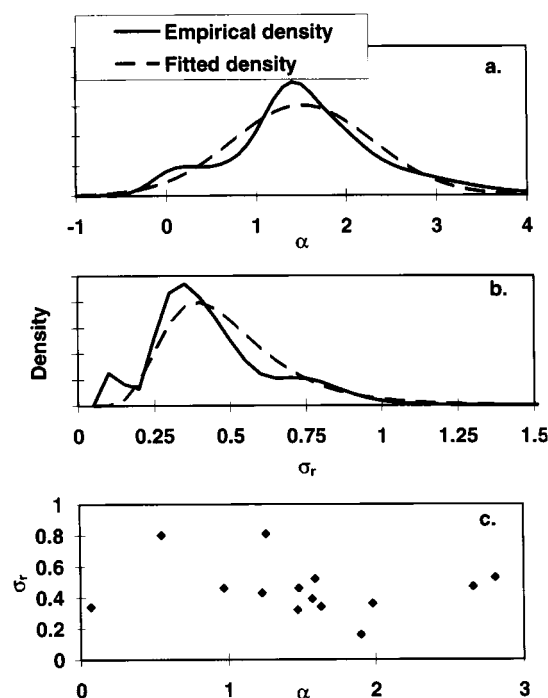
and a CV of 0.51 for the second series (Table 2). These prior pdfs were constructed by Monte Carlo simulation of the uncertainties in the following quantities: (i) the mean proportion of the population within the design in each year, (ii) the effective width swept by the net, as a fraction of the estimated width (this reflects the fraction of fish that encounter the net that are herded in from the trawl doors), (iii) the fraction that encounter the net that do not escape underneath the net, and (iv) the mean proportion of the fish detected as result of migration within the survey design and vessel avoidance (see Appendix 3 for details).

The priors for the selectivity parameters were diffuse but roughly centered about values consistent with previous as-

sessments. The priors for the slope at 50% selected for the commercial and the survey ogives were both Normal($1.5, 2^2$). The priors for age (in years) at 50% selected for the commercial and survey ogives were both Normal ($7.5, 3^2$). Both sets of priors were truncated at 0. It is generally improper to use the results of one assessment to specify priors in another. However, the priors chosen are very flat and do not affect the results (see below).

Importance function used for yellowfin sole

The importance function that was used to estimate the posterior for yellowfin sole was constructed using the following procedure. First, the parameters B_0, σ_r (when used), M, s_{50}^i, and q_i were log transformed so that when drawn from the multivariate t distribution, the untransformed values would be skewed as they are in the posterior (see Appendix 1 for details). Second, a nonlinear minimization procedure, based on a gradient search method that uses numerical derivatives (Press et al. 1992), was used to find the mode of the posterior. Third, the Hessian matrix was estimated using numerical derivatives of the log of the posterior (i.e., the likelihood times the prior) at its mode. However, only the priors for M and ϵ_y were included in estimating the Hessian matrix. By eliminating the other priors from the posterior, the estimated covariance is slightly larger than that for the posterior with all of the priors included. This procedure reduces the chance of using an importance function with sharper tails than the posterior and improves the efficiency of SIR. The importance function used was a multivariate t distribution with 25 degrees of freedom, the mean given by the estimate of the mode of the posterior (with all of the priors), and the variance given by the negative inverse of the Hessian matrix.

For the stock assessment, the resampling procedure selected 10 000 of the 3 000 000 draws from the importance function. The base case application of the stock assessment technique used the relative abundance indices in Table 1, catch-age data from the commercial fishery and the survey (McAllister 1995), the catch biomass series (Fig. 1; Table 1), the priors (Table 2), and the biological and technological parameters (Table 3).

Bayesian estimation of population dynamics model parameters and stock status

Bayesian estimates of uncertainty in model input parameters and the status of yellowfin sole in the eastern Bering Sea are conveyed by the marginal posterior pdfs of the following quantities: (i) the model input parameters (e.g., B_0, α, M, q_i, s_{50}^i, a_{50}^i, ϵ_y), (ii) current-year (1995) stock biomass (B_{1995}), and (iii) the amount of depletion from unexploited conditions (B_{1995}/B_0). Tests were conducted to assess the sensitivity of the posteriors to the use of age data, the use of a multivariate t distribution developed from properties at mode of the posterior as an estimate of the posterior distribution (this would lessen the lengthy Monte Carlo steps required by SIR (e.g., from 3 000 000 to 10 000)), and the amount of variability in recruitment, σ_r. We wanted to examine how well a multivariate t distribution approximated the posterior given by SIR. The multivariate t with 25 degrees of freedom approaches a multivariate normal.

In the first test, only the relative biomass data were used, and the age data excluded. The prior pdf was used as the

Table 3. Values of the model parameters that are assumed to be known without error in the Bayesian assessment of yellowfin sole.

Parameter	Value
L_∞ (cm)	35.8
κ (year^{-1})	0.147
t_0 (years)	0.47
δ_1 (g/cm^3)	0.009 721 7
δ_2	3.0564
Plus group in model: m (years)	21
Plus group in age data: m' (years)	18
$\mu_1-\mu_8$	0.000
μ_9	0.075
μ_{10}	0.135
μ_{11}	0.270
μ_{12}	0.375
μ_{13}	0.415
μ_{14}	0.460
μ_{15}	0.475
μ_{16}	0.486
μ_{17}	0.490
$\mu_{18}-\mu_m$	0.500

importance function and the selectivity parameters were set at the modal estimates from the baseline assessment. In the second test, the posterior from SIR was compared with a multivariate t with 25 degrees of freedom (eq. 4), parameter μ given by the posterior modal estimates of parameters (with M, B_0, q_i, and s_{50}^i on a log scale), and parameter V given by the Hessian estimate of covariance at the mode of the posterior. Distributions for derived variables such as current stock size were computed numerically by taking 10 000 random draws of θ from the multivariate t distribution, placing mass $1 / 10\ 000$ on each drawn θ_k, computing the derived variable, $g(\theta)$, for each θ_k, and then summing the mass in the appropriate interval over $g(\theta)$. In the third test, the value of σ_r was increased from 0.40 to the highest value plausible, 0.55 (i.e., the 99.7th percentile of the marginal posterior for σ_r, based on the Hessian estimate of the posterior variance in σ_r when σ_r is estimated jointly with the other parameters).

Results

Using catch-age data markedly updates the priors for most model input parameters (Figs. 3–7; Table 4). The marginal posteriors for most of the population dynamics model parameters have considerably lower variance than the priors (Table 4). For the estimated recruitment residuals in the last few years of the historic time series, the posterior reflects the priors because there are fewer data in the last years (Fig. 6). The posteriors for current stock size and depletion are also highly peaked (Fig. 7). The 95% Bayesian confidence intervals (CIs) about the recruited and survey stock biomass are very tight, especially for the earlier years when very large catches were taken and the population was heavily depleted (Figs. 1 and 8). The CIs become broader in the late 1970s, despite the addition of survey biomass and age data, mainly because the catches of yellowfin sole decreased

Fig. 3. Marginal posterior pdfs and priors of (a) α, (b) B_0, and (c) M. The marginal pdfs are shown for the baseline case, and the case in which the posterior was approximated by a multivariate t distribution with 25 degrees of freedom (see text for details).
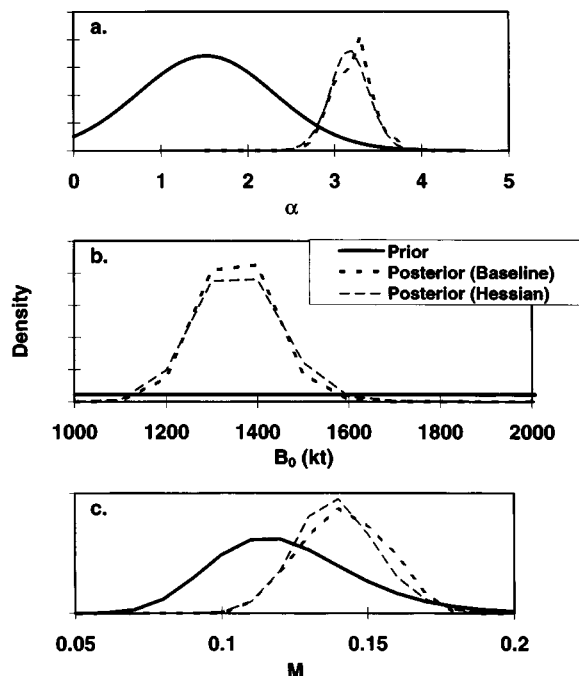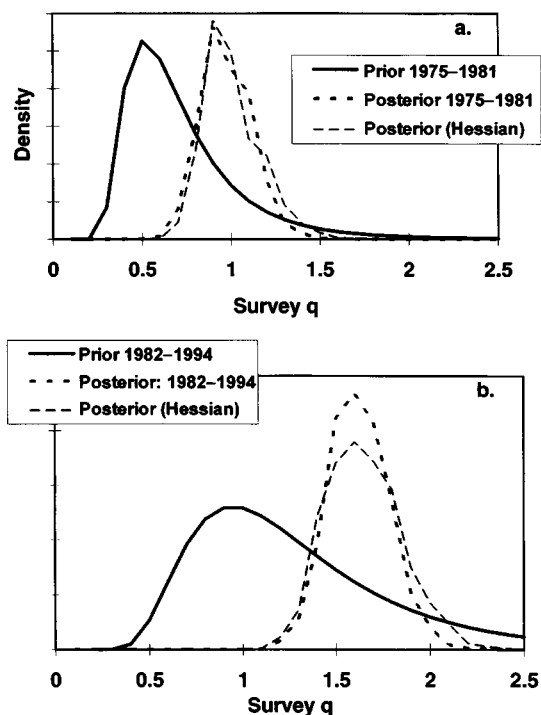
Fig. 5. Marginal posterior pdfs and priors of the (a) age at 50% selected and (b) slope of the logistic selectivity function at 50% selected for the fishery and the commercial catch. The marginals are shown for the baseline case only.
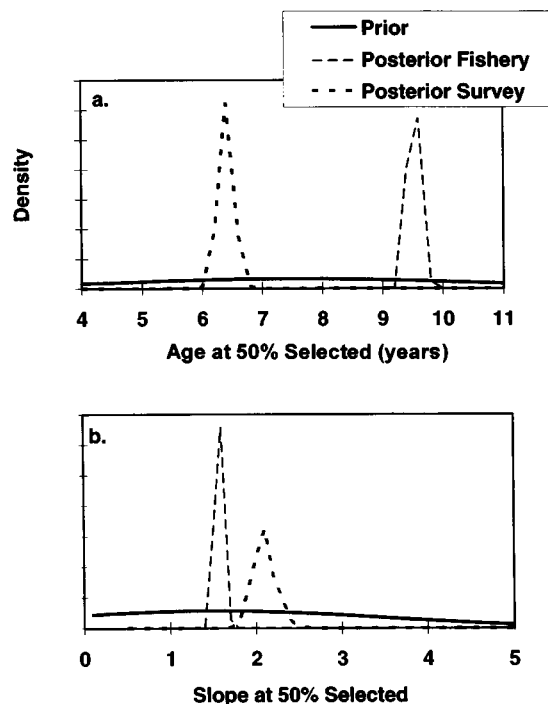
Fig. 4. Marginal posterior pdfs and priors of the catchability coefficient for the trawl survey for the periods (a) 1975–1981 and (b) 1982–1994. The marginals are shown for the baseline case and the case in which the posterior was approximated by a multivariate t distribution with 25 degrees of freedom (see text for details).

Fig. 6. Marginal posterior pdfs and priors of the recruitment residuals ϵ_y for the years (a) 1979–1983 and (b) 1986–1990. The marginals are shown for the baseline case only.
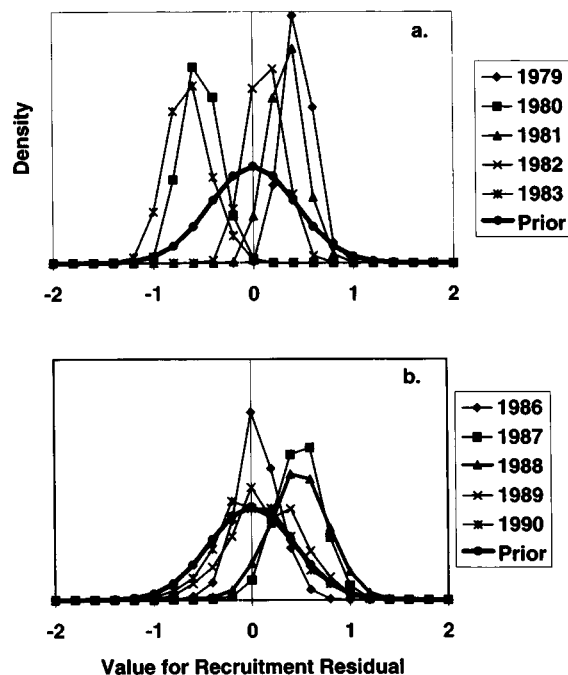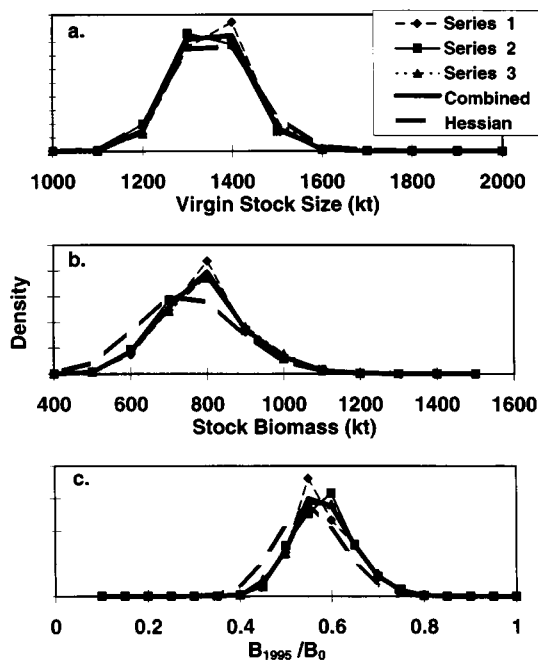
Table 4. Estimated ratio of the posterior variance to the prior variance for selected model input parameters.

Parameter	Baseline	Biomass data only	Age data only	$\sigma_r = 0.55$
B_0 ($\times 1000$ t)	0.005	0.507	0.006	0.006
α	0.082	0.550	0.078	0.111
M (year^{-1})	0.313	0.948	0.278	0.490
q_{75-81}	0.186	0.310	na	0.259
q_{82-94}	0.075	0.230	na	0.121
a_{50}^f (years)	0.001	na	0.001	0.001
a_{50}^s (years)	0.002	na	0.002	0.002
s_{50}^f	0.001	na	0.007	0.001
s_{50}^s	0.005	na	0.007	0.005
ε_{1985}	0.313	1.12	0.379	0.372

Fig. 7. Marginal posterior pdfs and priors of (a) virgin stock size (B_0), (b) biomass in 1995 (B_{1995}), and (c) stock depletion (B_{1995}/B_0). The marginals are shown for the baseline case with 3 million draws from the importance function and the three sequences of 1 million draws used and the case in which the posterior was approximated by a multivariaimates t distribution with 25 degrees of freedom (see text for details).

Fig. 8. Median and 95% confidence intervals for (a) recruited biomass and (b) surveyable biomass of yellowfin sole.

and inferences become less precise when fishing mortality rates are low (Bence et al. 1993). The plot of the modal estimates of recruitment versus spawning potential suggests a fairly well pronounced Ricker stock–recruit relationship with some of the largest recruitments occurring under the most depleted conditions (Fig. 9). However, there is still considerable scatter about the modal stock–recruit function as reflected by the modal recruitment CV (σ_r) of 0.40.

When only survey biomass data are used, the marginal posteriors for model input parameters are not much sharper than the priors and the marginal pdfs for quantities of interest such as current stock size are relatively flat (Figs. 10 and 11;
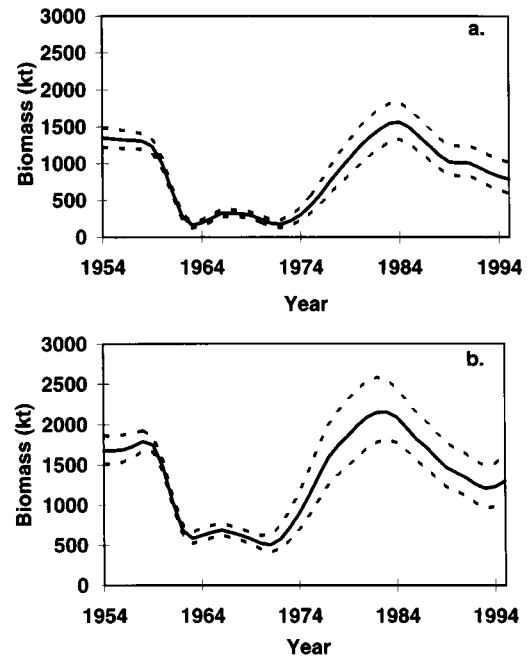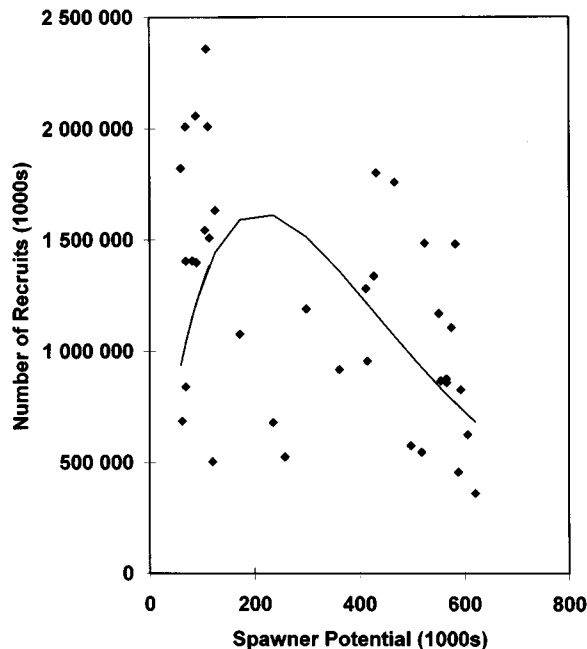
Table 4). The survey biomass data suggest that stock size is higher and the stock less productive than under the baseline scenario when both age and biomass data are used.

The sampling efficiency in estimating the posterior was much greater when the model was fit to biomass data only. For example, the maximum $w(\theta_k)$ for a single draw from the importance function was consistently less than 1% of the total at less than 500 000 draws. In contrast, the maximum $w(\theta_k)$ was 4% after 3 million draws from the importance function when age data were also included (the maximum $w(\theta_k)$ were 12, 4, and 3% in each consecutive sequence of 1 million in the 3 million draws) (Table 5). See Discussion for comments on sampling efficiency.

In test 2, we compared posteriors obtained by using the SIR algorithm (called the SIR estimator) with a multivariate t based on the posterior modal parameter values for the mean, the Hessian estimate of the covariance of the posterior at the

294

Can. J. Fish. Aquat. Sci. Vol. 54, 1997

Fig. 9. Modal estimates of the Ricker stock–recruit function and recruitment and spawner potential for the series 1954–1990.

Fig. 10. Marginal posterior pdfs and priors of (*a*) α, (*b*) B_0, and (*c*) *M*. The marginals are shown for the baseline case and the cases in which $\sigma_r = 0.55$, and only the survey biomass data are used.

mode, and degrees of freedom set at 25 (called the Hessian estimator, Figs. 3, 4, and 7). For most model input parameters, the Hessian estimator was reasonably close to the SIR estimator. However, the Hessian estimator favored smaller values for natural mortality rate (Fig. 3). For the derived quantities such as current stock size, the Hessian estimator was noticeably more pessimistic (Fig. 7). Current stock size is lower and depletion is more pronounced under the Hessian estimator. The differences in results between the two estimators appear to be due to real differences in the distributions rather than due to sampling variability. For example, the three replicate estimates of the posterior (of 1 million samples each) from SIR were much closer to each other than to the one estimated by the multivariate *t* (Fig. 7). Differences in marginals between the SIR estimator and the multivariate *t* do not necessarily imply that the posterior modal value for θ is different from that obtained by maximizing the posterior density function. The differences in estimated marginals between the posterior and the multivariate *t* could result from differences in their shape (e.g., the posterior could be banana-shaped, while the multivariate *t* is oval). Therefore, we conclude that at least for the yellowfin sole data analyzed here, the multivariate *t* with the posterior mode and Hessian estimate of covariance provides significantly poorer estimates of the posterior pdf relative to the SIR estimator.

The effect of replacing the value for σ_r of 0.40 with 0.55 was negligible: only slight changes in the marginal posterior pdfs of parameters and other quantities were observed. The variances in posterior pdf for model input parameters increased, as might be expected by increasing the amount of recruitment variability, but only slightly (e.g., the SD in B_0 under $\sigma_r = 0.55$ was 87 kt and under the baseline was 79 kt, see Table 4).

Fig. 11. Marginal posterior pdfs of (*a*) virgin stock size (B_0), (*b*) biomass in 1995 (B_{1995}), and (*c*) stock depletion (B_{1995}/B_0). The marginals are shown for the baseline case and the cases in which $\sigma_r = 0.55$, and only the survey biomass data are used.

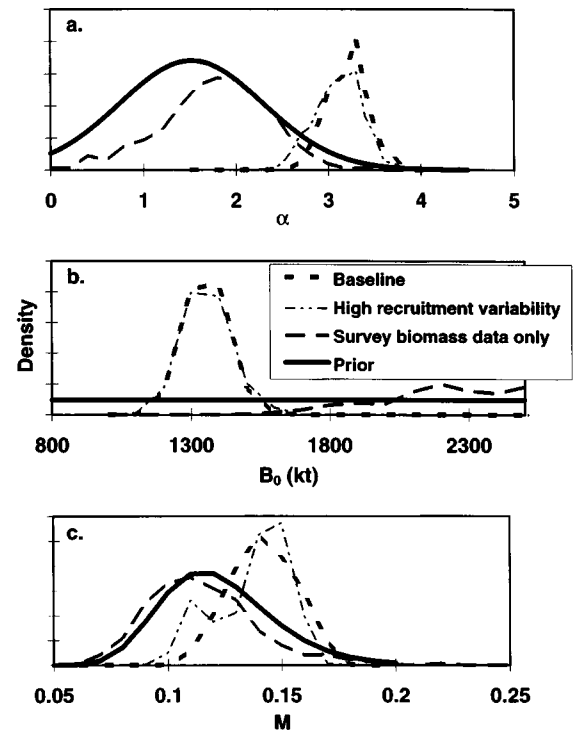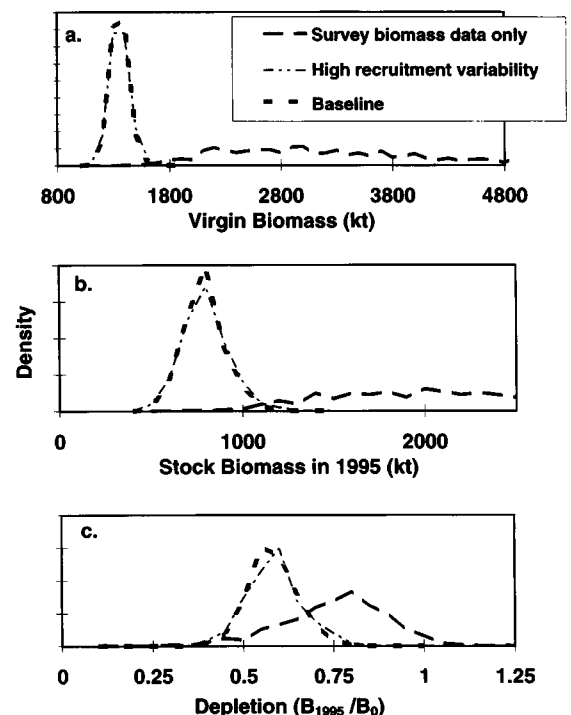**Table 5.** Posterior expectations and standard errors (in parentheses) for virgin biomass (B_0), current stock size (B_{1995}), and depletion (B_{1995}/B_0) for three replicate posteriors (each replicate was obtained by taking a series of 1 million i.i.d. draws from a multivariate t importance function; biomass values are in millions of tonnes).

Replicate	No. of draws	Max (% weight)	B_0	B_{1995}	B_{1995}/B_0
1	1 million	11.8	1.36 (0.08)	0.79 (0.10)	0.58 (0.06)
2	1 million	4.1	1.35 (0.08)	0.79 (0.11)	0.58 (0.07)
3	1 million	3.2	1.36 (0.08)	0.79 (0.11)	0.58 (0.07)
Combined	3 million	4.1	1.35 (0.09)	0.79 (0.11)	0.58 (0.07)

Discussion

This paper extends the use of the SIR algorithm and demonstrates that the use of catch-age data, in addition to relative abundance indices, for estimating a Bayesian posterior distribution can lead to markedly improved precision in estimates of population dynamics model parameters. The data series used in the example were highly informative because of the highly varied exploitation history of the population (i.e., high fishing effort followed by much lower effort, Hilborn and Walters 1992), the long time series of catch-age data, the existence of an independent survey of biomass and age frequency, and the high precision of the data series. For example, yellowfin sole is considered to be one of the most precisely aged fishes in the AFSC (Kimura and Lyons 1991).

However, the precision in posterior estimates is overestimated because several sources of uncertainty were ignored. These include uncertainty over the amount of recruitment variability (σ_r), the recruitment residuals, ϵ_y, for the earlier segment of the time series (1934–1945) (ages >8 years in 1954), mass at age, fecundity at age, catch biomass, the functional forms of survey and fishery selectivity and of the stock–recruit relationship, nonstationarity in model parameters, and autocorrelation in recruitment residuals. Furthermore, the multinomial likelihood function, often used in catch-age assessment methods (e.g., Methot 1990), may not adequately express the variability in the catch-age data. This is because sample units (e.g., trawl survey hauls) may capture groups of similarly aged individuals rather than individuals at random from the population, an assumption of the multinomial sampling process. However, in recent tests that we have conducted, estimated posterior variances were similar when the multivariate logistic likelihood function, derived specifically for catch-age data (Schnute and Richards 1995), replaced the multinomial. Even if additional uncertainties were effectively incorporated in a Bayesian assessment, the precision would still be considerably higher than that obtainable using abundance indices only.

In instances such as the current one, when data are highly informative, the issue of the choice of priors diminishes (Box and Tiao 1973). However, in the current procedure, informative priors are still needed for the amount of recruitment variability, σ_r, and the rate of natural mortality, M. These priors are necessary for the development of the importance function and to permit reasonably efficient SIR (McAllister et al. (1994).

Alternative Monte Carlo methods for estimating posteriors in Bayesian fisheries assessments include adaptive importance sampling (AIS) and Markov Chain Monte Carlo (MCMC) methods (Givens 1993a, 1993b; Kinas 1993; A.E. Punt, Commonwealth Scientific and Industrial Research Organization (CSIRO), Hobart, Tasmania, Australia, personal communication). In contrast with SIR, AIS is an iterative importance sampling procedure that often uses as an importance function a finite mixture of multivariate pdfs, such as the multivariate Student density (Kinas 1993; West 1993). In AIS, the normalized importance ratios are used to develop a new importance function. This process is repeated until the importance function converges on the posterior distribution. A potential advantage of AIS over SIR is that AIS may be computationally more efficient than SIR, especially for estimating complex (e.g., multimodal) posteriors (Kinas 1993). However, AIS can be more difficult to implement because AIS is considerably more complicated than SIR. Furthermore, it is unclear whether AIS could deal effectively with the large number of uncertain parameters in catch-age analysis.

In contrast, MCMC methods are based on iterative Markovian updating schemes for estimating posteriors. "Sampling" entails a random walk over the posterior probability surface. MCMC methods can be easier to implement than SIR because they may require less coding and initial start-up time than importance sampling methods. Also, estimation of an importance function is not required. However, MCMC methods may be computationally less efficient than importance sampling methods (Smith 1991; Givens 1993a). For example, when we applied the Hastings–Metropolis version of the MCMC algorithm (Metropolis et al. 1953; Hastings 1970; Gelfand and Smith 1990) to the data for yellowfin sole, it took several times longer than the SIR algorithm to obtain a reasonably precise estimate of the posterior (Hastings–Metropolis procedure provided by A.E. Punt). Reassuringly, the distribution estimated by Hastings–Metropolis was very similar to that provided by SIR (Fig. 12). Tuning MCMC so that the draws taken are independent, however, is not straightforward (Raftery and Lewis 1992). In addition, there are some conditions in which the MCMC may not necessarily converge on the posterior (e.g., when the posterior surface is multimodal or not log-concave, Newton and Raftery 1994). Furthermore, testing for convergence is also not straightforward. Because of these various difficulties with implementation, it may be prudent to apply both the SIR and MCMC methods whenever time permits to test whether results agree.

As a further advantage, SIR output is more concise and versatile than MCMC output. SIR output for each sampled

Fig. 12. Marginal posterior and prior pdfs of (*a*) α, (*b*) virgin stock size (B_0), and (*c*) stock biomass in 1995. The marginals are shown for the baseline case estimated by SIR and the Hastings–Metropolis algorithms (see Discussion).
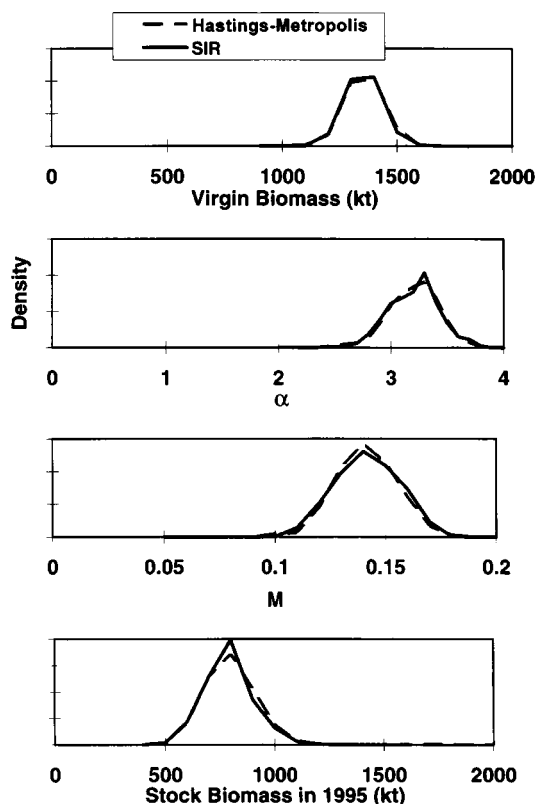
θ_k can be stored in two numbers: the initial random number seed that was used to generate each θ_k and the importance ratio, $w(\theta_k)$. These can easily be used to generate marginal posterior distributions of any quantity of interest in a stock assessment or decision analysis. Furthermore, it is unnecessary to rerun SIR to evaluate the sensitivity of functions of the posterior of θ (e.g., quantities from a decision analysis) to different priors. Instead, one simply divides each baseline $w(\theta_k)$ by the baseline prior $p(\theta_k)$ and multiplies by the new prior $p''(\theta_k)$ to give $w''(\theta_k)$. This avoids redoing the time-consuming steps of obtaining draws of θ_k from $h(\theta)$ and calculating the likelihood for each θ_k. The resampling step to obtain a set of draws from the new posterior of θ is then based on $w''(\theta_k)$ rather than $w(\theta_k)$.

In contrast, with MCMC, the full parameter vector must be stored for each sampled θ_k. Furthermore, to test the sensitivity of functions of the posterior of θ to different priors, it is necessary to rerun the MCMC procedure. This is because the random walk that produces the Markov Chain and the draws of θ from the posterior is a function of the likelihood and the prior. Therefore, the random walk must be reinitiated to obtain a set of draws of θ from a posterior that has a different prior. Despite these minor drawbacks, MCMC appears to be a promising alternative for Bayesian catch-age analysis. Further research on the application of MCMC (e.g., the Hastings–Metropolis algorithm) to catch-age data is required before an MCMC algorithm can be presented in detail as an alternative to SIR for Bayesian stock assessment and decision analysis.

When SIR is used, diagnostics are required to quickly evaluate whether the importance function chosen will provide a reasonably precise estimate of the posterior in an acceptable duration. They are also required to evaluate whether the estimated posterior is sufficiently close to the actual posterior. The diagnostics that we suggest are ad hoc but appear to work well in a variety of applications to catch-age data. In contrast, Kinas (1993) used the entropy relative to uniformity (ERU) which indicates how close the importance function is to the posterior. ERU is therefore not an absolute measure of sampling efficiency (i.e., it does not necessarily indicate whether an importance function will provide an acceptably precise estimate of the posterior or whether the final estimate of the posterior is sufficiently precise). For AIS, ERU is useful because a series of importance functions are used and the convergence of the importance function to the posterior can be observed. However, ERU is not appropriate for SIR because only one importance function is used. For example, there may be a large difference between the importance function and the posterior (i.e., ERU may be very low); but with many draws the final estimate of the posterior may be very precise.

We present the multivariate *t* distribution as an importance function that can be used for Bayesian catch-age analysis. However, in the example, this appears to be a relatively inefficient importance function because of the lengthy computing time, large number of draws used (3 million), and relatively high value for the maximum importance weight (i.e., 4%, 120 000 times higher than if the ideal importance function, the posterior, was used). This compares with 20 000 draws and a maximum weight <1% in McAllister et al. (1994) and about 300 000 draws and <1% in Raftery et al. (1995). Furthermore, it was not possible to include the prior for σ_r because SIR became too inefficient. Fixing σ_r at its modal value appeared to be a reasonable remedy and does not appear to bias the posterior (see McAllister et al. 1994). The SIR algorithm also performed poorly when all of the recruitment residuals at the beginning of the time series (the period during which no catch-age data exist) were treated as uncertain. Fixing the earliest residuals in the time series to zero remedied the situation and did not appear to bias the estimates of the posterior distribution.

Despite the inefficiencies of the multivariate *t*, an acceptably precise estimate of the posterior was obtained within a tolerable amount of computing time. For example, three replicate posteriors, each obtained by taking 1 million draws from the importance function, closely resemble each other (Fig. 7). Furthermore, the posterior expectations and standard errors for, e.g., B_0, B_{1995}, and B_{1995}/B_0 were very close despite the relatively large weights in each series (11.8, 4.1, and 3.2%) (Table 5). In other applications of the multivariate *t* distribution to catch-age data, much lower maximum weights have been obtained after fewer draws (e.g., <1% in <500 000 draws). Problems with efficiency could also be overcome by using better approximations of the posterior for the importance function or by using adaptive importance sampling (see Oh and Berger 1992; Kinas 1993; Newton and Raftery 1994).

Our approach differs in a variety of ways from that of Walters and Punt (1994), who used catch-age data to estimate a posterior distribution for the probability of sustaining current stock size from alternative allowable catch options.

For example, their estimation algorithm was based on virtual population analysis (backward calculation of numbers at age) tuned to a survey biomass series. Only few model parameters were treated as uncertain and a grid-based approach was used to compute the posterior distribution. Also, the posteriors on current-year stock biomass computed in Walters and Punt (1994) were more diffuse than the one in this paper because the time series of data was shorter and data were less informative.

Finally, the application of the procedure to yellowfin sole in the eastern Bering Sea demonstrates a simple way of incorporating uncertainty in scaling (via the q parameter) absolute survey abundance estimates. Clearly, uncertainty in this parameter can have broad implications for fishery management because the assessment of current stock size is the critical issue. For example, the estimates of population biomass given by the baseline Bayesian assessment were about 50% of those given with q fixed at 1.0 and assumed known with certainty. Also, by ignoring uncertainty in the value of q, the uncertainty in the estimate of current biomass will be underestimated. Hence, we advocate treating q as an estimated parameter and expressing stock assessment uncertainty using estimates of posterior probability distributions. Information about q (e.g., from studies on gear efficiency) that is not contained in the likelihood function can be included in the prior as has been done in the current paper. The idea that biases can result from treating q as known and fixing it at some value is not new but needs to be reemphasized because of existing practices.

The Bayesian estimation approach developed in the current paper is flexible and intuitively appealing and could be applied in a wide variety of situations involving catch-age data. For example, the procedure has been applied successfully to other sets of fishery age data in the Pacific Northwest (J. Ianelli), New Zealand (T. Hammond and R. Hilborn, School of Fisheries, University of Washington, Seattle, Wash.), England, and Australia (A.E. Punt, CSIRO, Hobart, Tasmania). The multivariate pdf that is estimated by the procedure is not only useful in stock assessment. It is also a key input to decision analyses of alternative management policies (Francis et al. 1992; McAllister et al. 1994). The posterior pdf of population parameters for yellowfin sole that we estimated is used in a companion paper (McAllister and Pikitch 1997) to evaluate the potential trade-offs among alternative trawl survey designs in the eastern Bering Sea.

Acknowledgments

References

Arnold, S.F. 1990. Mathematical statistics. Prentice Hall Inc., Englewood Cliffs, N.J.

Bence, J.R., Gordoa, A., and Hightower, J.E. 1993. Influence of age-selective surveys on stock synthesis assessments. Can. J. Fish. Aquat. Sci. **50**: 827–840.

Berger, J.O. 1985. Statistical decision theory and Bayesian analysis. 2nd ed. Springer-Verlag, New York.

Box, G.E.P., and Tiao, G.C. 1973. Bayesian inference in statistical analysis. Addison-Wesley, Reading, Mass.

Clark, W.G. 1991. Groundfish exploitation rates based on life history parameters. Can. J. Fish. Aquat. Sci. **48**: 734–750.

Engås, A., and Ona, E. 1990. Day and night fish distribution pattern in the net mouth area of the Norwegian bottom-sampling trawl. Rapp. P.-V. Reun. Cons. Int. Explor. Mer, **189**: 123–127.

Francis, R.I.C.C. 1992. Use of risk analysis to assess fishery management strategies: a case study using orange roughy (*Hoplostethus atlanticus*) on the Chatham Rise, New Zealand. Can. J. Fish. Aquat. Sci. **49**: 922–930.

Francis, R.I.C.C., Robertson, D.A., Clark, M.R., and Coburn, R.P. 1992. Assessment of the QMA 3B orange roughy fishery for the 1992/93 season. N.Z. Fish. Assess. Res. Doc. No. 92/4.

Gelfand, A.E., and Smith, A.F.M. 1990. Sampling-based approaches to calculating marginal densities. J. Am. Stat. Assoc. **85**: 398–409.

Geweke, J.F. 1989. Bayesian inference in econometric models using Monte Carlo integration. Econometrica, **57**: 1317–1340.

Givens, G.H. 1993*a*. Contribution to discussion on the meeting on 'The Gibbs sampler and other Markov chain Monte Carlo methods.' J. R. Stat. Soc. B, **55**: 75–76.

Givens, G.H. 1993*b*. A Bayesian framework and importance sampling methods for synthesizing multiple sources of evidence and uncertainties linked by a complex mechanistic model. Ph.D. thesis, Department of Statistics, University of Washington, Seattle, Wash.

Gunderson, D.R. 1993. Surveys of fisheries resources. John Wiley & Sons, Inc., New York.

Harden-Jones, F.R., Margetts, A.R., Greer Walker, M., and Arnold, G.P. 1977. The efficiency of the Granton otter trawl determined by sector-scanning sonar and acoustic transponding tags. Rapp. P.-V. Reun. Cons. Int. Explor. Mer, **170**: 45–51.

Hastings, W.K. 1970. Monte Carlo sampling methods using Markov chains and their applications. Biometrika, **57**: 97–109.

Hilborn, R., and Walters, C.J. 1992. Quantitative fisheries stock assessment: choice, dynamics, and uncertainty. Chapman and Hall, New York.

Hoff, R.Z. 1989. Spatial distributions and abundance trends for two species of flatfish in the eastern Bering Sea. M.S. thesis, University of Washington, Seattle, Wash.

Ianelli, J., and Heifetz, J. 1995. Decision analysis of alternative harvest policies for the Gulf of Alaska Pacific ocean perch fishery. Fish. Res. (Amsterdam), **24**: 35–63.

Kimura, D.K. and Lyons, J.J. 1991. Between-reader bias and variability in the age-determination process. Fish. Bull. **89**: 53–60.

Kinas, P.G. 1993. Bayesian statistics for fishery stock assessment and management. Ph.D. thesis, Department of Statistics, University of British Columbia, Vancouver, B.C.

Kinas, P.G. 1996. Bayesian fishery stock assessment and decision making using adaptive importance sampling. Can. J. Fish. Aquat. Sci. **53**: 414 423.

McAllister, M.K. 1995. Using decision analysis to choose a design for surveying fisheries resources. Ph.D. thesis, School of Fisheries, University of Washington, Seattle, Wash.

298

Can. J. Fish. Aquat. Sci. Vol. 54, 1997

McAllister, M.K., and Pikitch, E.K. 1997. A Bayesian approach to choosing a design for surveying fishery resources: application to the eastern Bering Sea trawl survey. Can. J. Fish. Aquat. Sci. **54**: 301–311.

McAllister, M.K., Pikitch, E.K., Punt, A.E., and Hilborn, R. 1994. A Bayesian approach to stock assessment and harvest decisions using the sampling/importance resampling algorithm. Can. J. Fish. Aquat. Sci. **51**: 2673–2687.

Methot, R.D. 1990. Synthesis model: an adaptable framework for analysis of diverse stock assessment data. *In* Proceedings of the Symposium on Applications of Stock Assessment Techniques to Gadids. *Edited by* L. Low. Int. North Pac. Fish. Comm. Bull. **50**: 259–277.

Metropolis, N., Rosenbluth, A.W., Rosenbluth, M.N., Teller, A.H., and Teller, E. 1953. Equations of state calculations by fast computing machines. J. Chem. Phys. **21**: 1087–1092.

Myers, R.A., Bridson, J., and Barrowman, N.J. 1993. Summary of worldwide stock and recruitment data. Can. Tech. Rep. Fish. Aquat. Sci. No. 2024.

Nash, J.C. 1990. Compact numerical methods for computers: linear algebra and function minimization. 2nd ed. Adam Higler Ltd., New York.

Newton, M.A., and Raftery, A.E. 1994. Approximate Bayesian inference with the weighted likelihood bootstrap. J. R. Stat. Soc. B, **56**: 3–48.

Oh, M.S., and Berger, J.O. 1992. Adaptive importance sampling in Monte Carlo integration. J. Stat. Comput. Simul. **41**: 143–168.

Pennington, M., and Vølstad, J.H. 1994. Assessing the effect of intra-haul correlation and variable density on estimates of population characteristics from marine surveys. Biometrics, **50**: 725–732.

Press, W.H., Teukolsky, S.A., Vetterling, W.T., and Flannery B.P. 1992. Numerical recipes in Fortran. The art of scientific computing. 2nd ed. Cambridge University Press, New York.

Punt, A.E. 1993. The implications of some multiple stock hypotheses for Chatham Rise orange roughy. N.Z. Fish. Assess. Res. Doc. No. 93/16.

Punt, A.E., Butterworth, D.S., and Penny, A.J. 1993. Stock assessment and risk analysis for the South Atlantic population of albacore (*Thunnus alalunga*). ICCAT Doc. No. SCRS/93/83. International Commission for the Conservation of Atlantic Tunas, La Jolla, Calif.

Raftery, A.E., and Lewis, S.M. 1992. How many iterations in the Gibbs sampler? *In* Bayesian Statistics 4: Proceedings of the Fourth Valencia International Meeting, April 15–20, 1991. *Edited by* J.M. Bernardo, J.O. Berger, P.A. David, and A.M. Smith. Clarendon Press, Oxford. pp. 763–773.

Raftery, A.E., Givens, G.H., and Zeh, J.E. 1995. Inference from a deterministic population dynamics model for bowhead whales. J. Am. Stat. Assoc. **90**: 402–416.

Ripley, B.D. 1987. Stochastic simulation. John Wiley & Sons, Inc., New York.

Rubin, D.B. 1987. Comment on 'The calculation of posterior distributions by data augmentation.' J. Am. Stat. Assoc. **82**: 543–546.

Rubin, D.B. 1988. Using the SIR algorithm to simulate posterior distributions. *In* Bayesian Statistics 3: Proceedings of the Third Valencia International Meeting, June 1–5, 1987. *Edited by* J.M. Bernardo, M.H. Degroot, D.V. Lindley, and A.M. Smith. Clarendon Press, Oxford. pp. 385–402.

Schnute, J.T., and Richards, L.J. 1995. The influence of error on population estimates from catch–age models. Can. J. Fish. Aquat. Sci. **52**: 2063–2077.

Smith, A.F.M. 1991. Bayesian computational methods. Philos. Trans. R. Soc. Lond. A Math. Phys. Sci. **337**: 369–386.

Stocker, M., Hand, C.M., and McAllister, M.K. 1994. Pacific cod. *In* Groundfish stock assessment for the west coast of Canada in 1993 and recommended yield options for 1994. *Edited by* M.

Stocker. Can. Tech. Rep. Fish. Aquat. Sci. No. 1975. pp. 50–92.

van Dijk, H.K., and Kloek, T. 1983. Monte Carlo analysis of skew posterior distributions: an illustrative economic example. Statistician, **32**: 216–233.

Walters, C.J., and Ludwig, D. 1994. Calculation of Bayes posterior probability distributions for key population parameters: a simplified approach. Can. J. Fish. Aquat. Sci. **51**: 713–722.

Walters, C.J., and Punt, A.E. 1994. Placing odds on sustainable catch using virtual population analysis and survey data. Can. J. Fish. Aquat. Sci. **51**: 946–958.

West, M. 1992. Modeling with mixtures. *In* Bayesian Statistics 4: Proceedings of the Fourth Valencia International Meeting, April 15–20, 1991. *Edited by* J.M. Bernardo, J.O. Berger, P.A. David, and A.M. Smith. Clarendon Press, Oxford. pp. 503–524.

West, M. 1993. Approximating posterior distributions with mixtures. J. R. Stat. Soc. B, **55**: 409–422.

Wilderbuer, T.K., Walters, G.E., and Bakala, R.G. 1992. Yellowfin sole, *Pleuronectes asper*, of the eastern Bering Sea: biological characteristics, history of exploitation, and management. Mar. Fish. Rev. **54**: 1–18.

Appendix 1

Derivation of the importance function when some variables in the multivariate t distribution are log transformed

Some of the population dynamics model parameters (e.g., M, B_0, σ_r, q_i, s_{50}^i in Table 2) were log transformed for use in the multivariate t importance function. This was to reproduce the skewness in the marginal posteriors for these parameters and improve the efficiency of SIR (van Dijk and Kloek 1983). During importance sampling, the importance ratios, $w(\theta_k)$, are computed:

$$(1.1) \qquad w(\theta_k) = \frac{L(\text{data} \mid \theta_k)p(\theta_k)}{h(\theta_k)}$$

where θ_k is one particular realization of the vector of population dynamics model input parameters θ, $w(\theta_k)$ is the importance ratio, $L(\text{data} \mid \theta_k)$ is the likelihood function of the data evaluated at θ_k, $p(\theta_k)$ is the prior pdf evaluated at θ_k, and $h(\theta_k)$ is the importance function evaluated at θ_k (Berger 1985, p. 263). The posterior of interest is for the vector of untransformed parameters, θ. The importance function, $h(\theta)$ in eq. 1.1, thus must be a density function of the untransformed variables (van Dijk and Kloek 1983).

With log transformation of some of the variables in the multivariate t distribution, the derivation of the resulting density function for the untransformed variables is fairly simple. It can be shown that if some of the variables in a multivariate t distribution are on a log scale, then the importance function for the untransformed variables is given by

$$(1.2) \qquad h(\theta \mid d, \mu, V) = \text{abs}[\det J]\frac{\Gamma((d+p)/2)}{(\det V)^{1/2}(\pi d)^{p/2}\Gamma(d/2)}$$

$$\times \left(1 + \frac{1}{d}(\theta'' - \mu)'V^{-1}(\theta'' - \mu)\right)^{-(d+p)/2}$$

where θ'' is the vector of estimated model input parameters, some of them log transformed, d is the degrees of freedom, μ is the mean of the vector θ'', V is a $p \times p$ positive definite matrix, and $\text{abs}[\det J]$ is the absolute value of the Jacobian,

the determinant of the matrix J whose (i,j)th component is (Arnold 1990)

$$J_{i,j} = \frac{\partial}{\partial \theta^j} z(\theta^i)$$

where $z(\theta^i)$ is the transformed parameter $\theta^{i\prime\prime}$ expressed as a function of the ith parameter θ^i.

When some of the members of the vector $\theta_k{}''$ are log transformed, J is a diagonal matrix with "1" for each of the untransformed parameters in $\theta_k{}''$ and "$1/\theta_k^j$" for each of the log transformed parameters θ_k^j where j is the index for parameter j. It can be shown that abs[det J] is simply the absolute value of the product of the terms in the diagonal of J:

$$(1.3) \quad \text{abs[det } J] = \text{abs}\left[\prod_{j=1}^{p}\left(\frac{1}{\theta_k^j}\right)^{(1-I_j)}\right]$$

where θ_k^j is the untransformed value for the jth parameter of the vector θ_k, $I_j = 0$ if θ_k^j is log transformed in $\theta_k{}''$ and $I_j = 1$ if not.

Note that the result in eq. 1.2 is equivalent to using an importance ratio in the transformed parameters. For example, the prior in the transformed parameters is given by

$$(1.4) \quad p(\theta'') = \frac{p(\theta)}{\text{abs[det } J]}.$$

This follows from the standard formula for a change of variables (Arnold 1990).

In the SIR algorithm, values for the vector θ'', some of them on a log scale, are drawn from a multivariate t distribution. This is equivalent to drawing θ_k from the density function in eq. 1.2. The density of the importance function $h(\theta)$ is evaluated at θ_k using eq. 1.2. The calculated density is then used to compute the importance ratio $w(\theta_k)$ (eq. 1.1).

Appendix 2

Likelihood function for catch-age data
The numbers of fish at age were considered to be multinomially distributed. An appropriate likelihood function for these data could be

$$(2.1) \quad \lambda^i = \sum_{j=1}^{n_i}\sum_{a=1}^{m'} x_{a,j}^i \log(c_{a,j}^i)$$

where $x_{a,j}^i$ is the observed number of fish, $c_{a,j}^i$ is the predicted proportion of fish at age a in the jth year of series i, n_i is the number of years in series i, and m' is the plus group for the age data (18 years for both series). However, owing to the sampling methodology, the assumption of multinomial sampling does not hold. The effective sample size per year, i.e., the number of independent sample units, is believed to be much smaller than the actual number of fish aged per year. This is because the fish are not sampled one by one at random from the entire population. Instead the fish tend to range about in groups of similar ages. Thus, when they are sampled, they are scooped up in nets along with fish of similar ages. It is therefore common practice not to use the original number aged as the effective sample size. Instead,

the proportions at age from the original sample are calculated by

$$(2.2) \quad o_{a,j}^i = \frac{x_{a,j}^i}{T_{i,j}}$$

where $T_{i,j}$ is number of fish aged in year j and series i, and a much smaller number is used for the sample size in each year ($T_{i,j}' < T_{i,j}$). It is also common practice to assume that the sample size is the same in each year (e.g., Methot 1990). Hence, T_i' is used instead of $T_{i,j}'$. The effective number of fish observed in each age a, year j, and series i is given by

$$(2.3) \quad x_{a,j}^i = T_i' o_{a,j}^i.$$

Therefore, the contribution of series i to the log likelihood is

$$(2.4) \quad \lambda^i = \hat{T}_i \sum_{j=1}^{n_i}\sum_{a=1}^{m'} o_{a,j}^i \log(c_{a,j}^i)$$

where $\hat{T}_i$ is the estimated effective sample size for the total number of fish sampled in each year. T_i was estimated by starting with a first guess at T_i and then estimating T_i by using (see below for derivation)

$$(2.5) \quad \hat{T}_{i,j} = \frac{\displaystyle\sum_{a=1}^{m'} c_{a,j}^i(1 - c_{a,j}^i)}{\displaystyle\sum_{a=1}^{m'}(o_{a,j}^i - c_{a,j}^i)^2}.$$

The values of $c_{a,j}^i$ that were used were obtained from the modal estimate from the posterior pdf of the model input parameters. The estimated effective sample size for series i was obtained from the mean of the estimates of effective sample size over the series

$$(2.6) \quad \hat{T}^i = \frac{1}{n_i}\sum_{j=1}^{n_i}\hat{T}_{i,j}.$$

The estimate, $\hat{T}_i$, was estimated iteratively by inputting the last estimate of $\hat{T}_i$ into the estimation procedure to obtain the mode of the posterior of model input parameters and then reestimating $\hat{T}_i$. The estimates of $\hat{T}_i$ converged at 104 and 102 for the trawl survey and catch age series, respectively.

Equation 2.5 is derived from the following:

$$(2.7) \quad \text{var}(o_{a,j}^i) = E(o_{a,j}^i - \bar{c}_{a,j}^i)^2 = \frac{\bar{c}_{a,j}^i(1 - \bar{c}_{a,j}^i)}{T_{i,j}}$$

where $\bar{c}_{a,j}^i$ is the true proportion of fish at age in series i in year j, assuming that the sample from the population is multinomially distributed. Using a normal approximation and eq. 2.7, the variance in the observed proportion can be approximated by

$$(2.8) \quad \text{var}(o_{a,j}^i) \cong (o_{a,j}^i - c_{a,j}^i)^2 \cong \frac{c_{a,j}^i(1 - c_{a,j}^i)}{T_{i,j}}$$

where $c_{a,j}^i$ is the posterior modal estimate of $c_{a,j}^i$. The estimate

of $T_{i,j}$ in eq. 2.5 is obtained by rearranging the following equality:

$$(2.9) \quad \sum_{a=1}^{m'} \text{var}(o_{a,j}^i) \cong \sum_{a=1}^{m'} (o_{a,j}^i - c_{a,j}^i)^2 \cong \sum_{a=1}^{m'} \frac{c_{a,j}^i(1 - c_{a,j}^i)}{T_{i,j}}.$$

Appendix 3

Constructing priors for catchability in trawl survey biomass indices

The catchability coefficients, q_i, relate the trawl survey abundance indices to the biomass of yellowfin sole in the eastern Bering Sea that is potentially vulnerable to trawl survey gear. Separate prior distributions for q_i were constructed for the two trawl series, one from 1975 to 1981 and the other from 1982 to 1994. This was because the trawl survey gear was modified in 1982 to make it more efficient for bottom dwelling species (e.g., starting in 1982, two extra 61-cm sections of chain were attached to the footrope of the net so that better bottom contact was achieved, Gunderson 1993).

The method used to construct the priors for q_i is analogous to the Monte Carlo approach used by McAllister et al. (1994) to construct a prior for a constant of proportionality for an acoustic biomass series. The procedure has five steps: (i) identify the variables that affect q_i, (ii) construct q_i as an explicit function of these variables, (iii) construct pdfs for these variables using the scientific literature, data that are not in the likelihood function in Bayesian posterior (Equation 1), and by consulting technical experts on the processes of interest, (iv) estimate a distribution for q_i using Monte Carlo simulation, and (v) fit a parametric density function to the estimated distribution.

The mean size of q_i for each time series depends on several factors. If values for each of these factors are assumed to be independent and are randomly generated from their respective distributions, the value for the catchability coefficient for the trawl series from 1982 to 1994 can be calculated using the equation

$$(3.1) \quad q_2 = p \cdot m'' \cdot c_2 \cdot h.$$

The value for q_1 can be calculated using

$$(3.2) \quad q_1 = p \cdot m'' \cdot \frac{c_2}{r} \cdot h.$$

These factors and their distributions are explained below.

Herding (h) is the ratio of the total number of fish that encounter the net to the number that encounter the net that are immediately before the net as it approaches. This factor reflects the bias introduced by having fish herded by the trawl doors into the estimated path swept by the net. The estimated path swept in the eastern Bering Sea trawl survey is that between the wing tips of the net. The trawl doors are large metal plates that drag along the bottom and help to keep the net spread wide open. Several studies indicate that these

increase the catching power of the net by scaring fish from outside into the path swept by the net (Harden Jones et al. 1977). David Somerton (AFSC, NMFS, Seattle, Wash.) has conducted some trawl experiments with trawl gear similar to that used in the eastern Bering Sea trawl survey to estimate this herding component. His studies and other studies that he is familiar with suggest the following distribution for h for flatfishes: h is a truncated Normal(2.25, 0.50^2) with $h > 0$.

Efficiency of the trawl net given that fish have encountered the net (c_i) for each trawl series i: this is the proportion of the fish that encounter the net that go into it and do not escape. Craig Rose (AFSC, NMFS, Seattle, Wash.) has conducted some underwater video experiments using the same trawl gear as that used in the eastern Bering Sea trawl survey and estimated that very few flatfish that encounter the net escape. This suggests the following distribution for c for the net used between 1982 and 1994: c_2 is a truncated Normal(0.90, 0.05^2) with $c_2 < 1$. Studies that compare the relative efficiency of the more recent gear with that of the gear used before 1982 suggest that the more recent gear may be two or more times more efficient than the older gear (Hoff 1989). The assumed distribution for this ratio, r, for several ranges of sizes of yellowfin sole in Hoff (1989) suggests the following distribution for r : r is lognormal(2.5, 0.35^2). The mean and the variance used are calculated from the means of the estimates and their variances in Hoff (1989) for size groups 20 cm and larger. The value of c_1 for the series before 1982 is given by c_2/r.

Proportion of the population of yellowfin sole in the eastern Bering Sea that on average across years is within the area of the survey (p): this variable is modeled as a function of the proportion of the biomass that is to the north of the design, P_N, and the proportion of the biomass that is to the east of the design, P_E. Work in McAllister (1995, Chap. 1) suggests that these proportions can be modeled as follows: $P_N \sim \text{Uniform}(0.05, 0.20)$ and $P_E \sim \text{Uniform}(0.05, 0.20)$. The proportion of the biomass within the design is given by

$$(3.3) \quad p = \frac{1}{1 + P_N + P_E}.$$

Bias from migration of fish within the design and vessel avoidance (Engås and Ona 1990) (m''): work in McAllister (1995, Chap. 2) suggests that if yellowfin sole migrate inshore at a rate of about 0.1 m/s, the expected bias is about -30%. At a migration rate of 0.2 m/s, this becomes -50%. At 0.5 m/s, this would be -75%. Vessel avoidance would also increase the negative bias in trawl biomass estimates. m'' was assumed to have a truncated Normal(0.60, 0.20^2) distribution with $m'' > 0$. The variance is a reflection of the amount of uncertainty in the interannual mean value of this quantity.

Any values for q_i that were less than 0.05 were discarded. Lognormal distributions were fitted to the resulting distributions for q_1 and q_2. The resulting distributions were $q_1 \sim$ lognormal(0.44, 0.59^2) and $q_2 \sim$ lognormal(0.97, 0.51^2).